Multi-omics Quantification of Species Variation of Escherichia coli Links Molecular Features with Strain Phenotypes

Escherichia coli strains are widely used in academic research and biotechnology. New technologies for quantifying strain-specific differences and their underlying contributing factors promise greater understanding of how these differences significantly impact physiology, synthetic biology, metabolic engineering, and process design. Here, we quantified strain-specific differences in seven widely used strains of E. coli (BL21, C, Crooks, DH5a, K-12 MG1655, K-12 W3110, and W) using genomics, phenomics, transcriptomics, and genome-scale modeling. Metabolic physiology and gene expression varied widely with downstream implications for productivity, product yield, and titer. These differences could be linked to differential regulatory structure. Analyzing high-flux reactions and expression of encoding genes resulted in a correlated and quantitative link between these sets, with strain-specific caveats. Integrated modeling revealed that certain strains are better suited to produce given compounds or express desired constructs considering native expression states of pathways that enable high-production phenotypes. This study yields a framework for quantitatively comparing strains in a species with implications for strain selection.